

①9 BUNDESREPUBLIK
DEUTSCHLAND



DEUTSCHES
PATENTAMT

⑫ Patentschrift
⑪ DE 3440613 C1

②1 Aktenzeichen: P 34 40 613.1-35
②2 Anmeldetag: 7. 11. 84
④3 Offenlegungstag: —
④5 Veröffentlichungstag
der Patenterteilung: 10. 4. 86

⑤1 Int. Cl. 4:
H 04 B 1/66
H 04 H 1/00
H 04 B 14/04
H 04 J 3/00

Behördeneigenthum

DE 3440613 C1

Innerhalb von 3 Monaten nach Veröffentlichung der Erteilung kann Einspruch erhoben werden

⑦3 Patentinhaber:

Institut für Rundfunktechnik GmbH, 8000 München,
DE

⑦4 Vertreter:

Konle, T., Dipl.-Ing., Pat.-Anw., 8000 München

⑦2 Erfinder:

Theile, Günther, Dr.-Ing., 8191 Thanning, DE

⑤6 Im Prüfungsverfahren entgegengehaltene
Druckschriften nach § 44 PatG:

US-Z.: The Bell System Journal, Okt. 1976,
S. 1069-1085;

⑤4 Verfahren zum digitalen Übertragen eines Rundfunk-Programmsignals

Zum digitalen Übertragen eines Rundfunk-Programmsignals wird das Frequenzband des Rundfunk-Programmsignals in Teilbänder aufgespalten. Die einzelnen Teilbänder werden unterschiedlich quantisiert und anschließend im Zeitmultiplex übertragen. Dabei erfolgt die Quantisierung jedes Teilbandes nach Maßgabe der für einen vorgegebenen Abhörlautstärkepegel geltenden Ruhehörschwelle und/oder nach Maßgabe der für die jeweilige Bandbreite des Teilbandes geltenden Mithörschwelle und/oder nach Maßgabe der für die jeweilige Frequenzlage des Teilbandes geltenden Amplitudenstatistik des Rundfunk-Programmsignals.

DE 3440613 C1

Patentansprüche:

1. Verfahren zum digitalen Übertragen eines Rundfunk-Programmsignals, bei dem das Frequenzband des Rundfunk-Programmsignals in Teilbänder aufgespalten wird sowie die einzelnen Teilbänder unterschiedlich quantisiert und anschließend im Zeitmultiplex übertragen werden, gekennzeichnet dadurch, daß die Quantisierung jedes Teilbandes durch folgende einzelne oder in Kombination angewendete Maßnahmen erfolgt:
 - a) nach Maßgabe der für einen vorgegebenen Abhörlautstärkepegel geltenden Ruhehörschwelle;
 - b) nach Maßgabe der für die jeweilige Bandbreite des Teilbandes geltenden Mithörschwelle;
 - c) nach Maßgabe der für die jeweilige Frequenzlage des Teilbandes geltenden Amplitudenstatistik des Rundfunk-Programmsignals erfolgt.
2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, daß die Quantisierung für jedes Teilband derart gewählt wird, daß der resultierende Quantisierungsrauschpegel ungefähr gleich dem für das jeweilige Teilband geltenden Wert der Ruhehörschwelle bei dem vorgegebenen Abhörlautstärkepegel ist.
3. Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, daß jedes Teilband komprimiert wird und die Kennlinie für diese Komprimierung derart gewählt wird, daß der daraus resultierende Quantisierungsrauschpegel ungefähr gleich dem an der jeweiligen unteren Teilbandgrenzfrequenz vorhandenen Wert der für die jeweilige obere Teilbandgrenzfrequenz geltenden Mithörschwellkurve liegt.
4. Verfahren nach einem der Ansprüche 1 bis 3, wobei das digitalisierte Rundfunk-Programmsignal empfangenseitig digital/analog-gewandelt wird, dadurch gekennzeichnet, daß die Teilbänder vor der Quantisierung unterschiedlich vorverstärkt werden und nach der empfangenseitigen Digital/Analog-Wandlung die Verstärkungsunterschiede ausgeglichen werden.
5. Verfahren nach einem der Ansprüche 1 bis 3, wobei das digitalisierte Rundfunk-Programmsignal empfangenseitig digital/analog-gewandelt wird, dadurch gekennzeichnet, daß ein Teilband vor der Quantisierung nach Maßgabe des komplementären Frequenzverlaufs der Ruhehörschwelle vorverzerrt wird und daß nach der empfangenseitigen Digital/Analog-Wandlung die Vorverzerrung ausgeglichen wird.
6. Verfahren nach Anspruch 4 oder 5, dadurch gekennzeichnet, daß der Grad der Vorverstärkung und/oder der Vorverzerrung von der Amplitude des jeweiligen Teilbandes gesteuert wird.

Die Erfindung geht aus von einem Verfahren nach dem Oberbegriff des Patentanspruchs 1. Ein solches Verfahren ist aus der Zeitschrift »The Bell System Journal«, Oktober 1976, Seiten 1069 bis 1085 bekannt.

Hochwertige Rundfunk-Programmsignale mit Frequenzspektren bis zu 15 kHz werden derzeit nur auf Kabel- und Richtfunkstrecken digital übertragen. Es gibt jedoch Überlegungen, im UKW-Bereich künftig eine digitale Hörfunkübertragung vorzusehen, um auch im Hörfunk die Qualität von digitalen Tonträgern zu erzielen.

Zur digitalen Übertragung von Sprachsignalen auf PCM-Strecken ist es bekannt (»The Bell System Technical Journal«, Oktober 1976, Seiten 1069 bis 1085), das Frequenzband des analogen Sprachsignals durch eine entsprechende Anzahl von Bandpässen in Teilbänder mit ansteigender Bandbreite aufzuspalten und jedes Teilband zu quantisieren, wobei die Abtastfrequenz für jedes Teilband mindestens der doppelten Bandbreite des betreffenden Teilbandes entspricht. Hierdurch ergeben sich unterschiedliche Abtastfrequenzen für die einzelnen Teilbänder. Desweiteren erfolgt für die Teilbänder eine zum Teil unterschiedliche Quantisierung, so daß sich insgesamt entweder bei vorgegebener Übertragungs-Bitrate für die anschließende Zeitmultiplex-Übertragung der einzelnen Teilbilder die Verteilung des Quantisierungsrauschens im Sinne einer Verbesserung der Signalqualität beeinflussen läßt oder bei vorgegebener Signalqualität die Übertragungs-Bitrate verringern läßt.

Die Aufgabe der Erfindung besteht darin, daß für die Übertragung von Sprachsignalen vorgesehene, bekannte Verfahren für die Übertragung von hochwertigen Rundfunk-Programmsignalen, insbesondere im UKW-Bereich, anzupassen, um auf diese Weise eine Verringerung der Übertragungs-Bitrate ohne Verschlechterung der Signalqualität zu ermöglichen.

Diese Aufgabe wird erfindungsgemäß durch die im kennzeichnenden Teil des Patentanspruchs 1 angegebenen Merkmale gelöst.

Vorteilhafte Ausgestaltungen des erfindungsgemäßen Verfahrens ergeben sich aus den Unteransprüchen.

Die Erfindung wird nachstehend anhand der Figuren 1 bis 10 näher erläutert.

Der bei einer PCM-Strecke mindestens zu übertragende Informationsfluß hängt bei fester Abtastfrequenz nur von der Zahl der Quantisierungsstufen ab, die je Abtastwert (sample) mindestens durchschnittlich übertragen werden müssen, um eine ausreichende Systemdynamik zu erhalten. Die Systemdynamik ist ausreichend, wenn der Quantisierungsfehler (das Quantisierungsrauschen) der PCM-Strecke gerade nicht hörbar ist. Bei der Wiedergabe muß also der Schalldruckpegel des Quantisierungsrauschens stets unterhalb der Ruhehörschwelle bzw. der Mithörschwelle des menschlichen Gehörs liegen.

Anpassung an die Hörschwellen

Diese Forderung führt bei den bekannten PCM-Systemen zu unnötig hohen Informationsflüssen, weil der »Breitband-PCM-Kanal« grundsätzlich nicht vollständig an den »frequenzgruppenbreit codierten Kanal« des Gehörs anpaßbar ist: Breitband-PCM-Systeme erzeugen ein breitbandiges Quantisierungsrauschen, dessen

Spektrum sich zur stark frequenzabhängigen Empfindlichkeit des Gehörs ungünstig verhält. Dies wird anhand von Fig. 1 veranschaulicht, wo der Frequenzverlauf der Ruhehörschwelle sowie einer Mithörschwelle des Gehörs eingetragen sind.

Der schraffierte Bereich in Fig. 1 unterhalb der Ruhehörschwelle entspricht dem gerade nicht wahrnehmbaren Schalldruckpegel. Das Spektrum des Leerkanalrauschens einer PCM-Strecke darf also den Frequenzverlauf der Ruhehörschwelle aufweisen, um unhörbar zu sein. Im Bereich hoher und tiefer Frequenzen darf die Energie des Leerkanalrauschens entsprechend höher sein als im Frequenzbereich 2 ... 5 kHz.

Dagegen erzeugen Breitband-PCM-Strecken näherungsweise ein weißes Leerkanalrauschen. Die Rauschleistung der Breitband-PCM-Strecke muß daher im gesamten Frequenzbereich so klein gehalten werden, daß die Ruhehörschwelle auch im empfindlichen Frequenzbereich 2 ... 5 kHz nicht überschritten wird.

Eine PCM-Strecke, deren Leerkanalrauschen weiß ist, überträgt somit einen irrelevanten Informationsfluß. Er ergibt sich aus der schraffierten Fläche unterhalb der Ruhehörschwelle und beträgt etwa 64 kbit/s. Der Informationsfluß einer PCM-Strecke läßt sich also ohne Qualitätsverlust um 64 kbit/s reduzieren, wenn das Spektrum des Leerkanalrauschens durch geeignete Codierung dem Verlauf der Ruhehörschwelle angepaßt wird.

Weiterhin ist eine breitbandige digitale Kompondierung nicht optimal bezüglich der Verdeckungseigenschaften des menschlichen Gehörs. Beispiele digitaler Kompondierungsprinzipien sind:

- Anwendung nichtlinearer Quantisierungsskalen, Momentanwertkompondierung;
- Anwendung mehrerer linearer Quantisierungsskalen mit unterschiedlicher Stufenhöhe und Übertragung des Skalenfaktors. Blockkompondierung (ein Skalenfaktor je Abtastwerte-Block, z. B. 32 samples).

Gegenüber linearer Quantisierung erzielen bekannte Verfahren eine Reduktion des Informationsflusses von etwa 100 ... 150 kbit/s, je nach Kompondierungsgrad. Dafür wird aber für hohe Signalpegel eine höhere Rauschleistung in Kauf genommen (»Kompondierungsrauschen«), d. h., das Quantisierungsrauschen des leeren Kanals steigt bei höheren Signalpegeln. Die mit höheren Signalpegeln verbundenen Verdeckungseffekte durch das Gehör reichen jedoch für kritische Programmarten nicht aus. Alle bisher vorgeschlagenen digitalen Kompondierungsverfahren (z. B. 14 auf 11 oder 14 auf 10) erzeugen ein u. U. deutlich wahrnehmbares Kompondierungsrauschen (»Rauschfahne«), und zwar auch bei Einsatz weitgehender Pre-Deemphasis (z. B. entsprechend CCITT Rec. J 17).

Breitbandige digitale Kompondierungen erzielen ohne Qualitätsverlust nur eine geringe Bitratenreduzierung, weil das breitbandige Kompondierungsrauschen die Mithörschwelle des Gehörs sehr leicht überschreiten kann. Fig. 1 zeigt die Mithörschwelle beispielsweise bei Übertragung eines Schmalbandrauschens ($f_m = 800$ Hz), und Wiedergabe mit 50 dB Schalldruck: die Empfindlichkeit des Ohres ist im Frequenzbereich oberhalb 2 kHz kaum geringer als ohne Maskierer. Das gilt sogar für Tiefpaßrauschen, $f_{gr} = 800$ Hz. Eine im Bereich 2 ... 5 kHz mangelhafte Verdeckung des Kompondierungsrauschens durch das Nutzsignal schränkt den möglichen Kompondierungsgrad stark ein. Damit ergeben sich aus den Mithörschwellenkurven des Gehörs weitere Möglichkeiten zur Reduktion des Informationsflusses: indem man das Spektrum des Kompondierungsrauschens durch geeignete Codierung dem Verlauf der Mithörschwellenkurven anpaßt, läßt sich der Kompondierungsgrad ohne Qualitätsverlust erheblich steigern. Dabei läßt sich in gleicher Weise wie bei der Ruhehörschwellenanpassung die erzielbare Einsparung an Kanalkapazität aus der Fläche unterhalb der jeweils wirksamen Mithörschwellenkurve ableiten.

Hörschwellenanpassung durch frequenzabhängige Quantisierung

Die Hörschwellen lassen sich auffassen als die – frequenzabhängigen – Pegel eines gehörinternen Rauschens. Wie in einem PCM-Kanal kann dabei ein Ruherauschen (Leerkanalrauschen) unterschieden werden vom signalabhängigen Rauschen. Die optimale Anpassung eines Tonübertragungskanal an das Gehör wäre daher die Nachbildung der gehörinternen Codierung, was nicht realisierbar ist. Es müssen vielmehr »künstliche«, nachrichtentechnisch realisierbare Codierungen benutzt werden, die das Rauschverhalten des Gehörs im Sinne einer minimalen Kanalkapazität nutzen können. Dies können Codierungen im Frequenz- oder Zeitbereich sein, feste oder adaptive Codierungen mit oder ohne »Gedächtnis«: sie können nur optimal sein, wenn das Rauschspektrum den Hörschwellen entspricht.

Hinsichtlich der entsprechenden Nutzung der Mithörschwellen ist zu beachten, daß das Gehör die Mithörschwellen innerhalb begrenzter Frequenzbänder (Frequenzgruppen) bildet. Die Zerlegung des Frequenzspektrums in Frequenzbänder führt offenbar auf die Verläufe der Mithörschwellenkurven. Nimmt man an, daß im Gehör ein signalabhängiges »Codierungsrauschen« auftritt, so läßt sich der Verlauf der Mithörschwellenkurve im Prinzip durch die frequenzgruppenbreite Bandbegrenzung des Codierungsrauschens erklären. Dabei läßt sich der gegenüber den Frequenzgruppenfiltern flachere Verlauf der Flanken, besonders der oberen Flanken bei hohen Pegeln, durch »Übersprechen« in die Nachbarfrequenzgruppen erklären.

Hieraus läßt sich das Konzept einer optimalen Codierung ableiten: Das Codierungsrauschen eines Kanals kann optimal durch das Nutzsignal verdeckt werden, wenn die Codierung nicht breitbandig, sondern frequenzgruppenbreit erfolgt. Insbesondere kann innerhalb jedes (beliebig schmalen) Bandpasses die Codierung beliebig genau an das Rauschverhalten des Gehörs angepaßt werden.

Die vorgenannten Erfindungsprinzipien sollen nunmehr anhand von PCM-Strecken betrachtet werden. Obwohl das Quantisierungsrauschen in einer PCM-Strecke ein sehr flaches Frequenzspektrum aufweist, kann durch Bandpaß-spezifische Quantisierung bzw. Kompondierung die Rauschanpassung erfolgen. Wie noch gezeigt werden soll, erweist sich diese sogenannte »Bandpaß-PCM« als wirkungsvoll für die Berücksichtigung sowohl der Ruhehörschwelle als auch der Mithörschwelle. Nimmt man an, daß die Ruhehörschwelle sich durch

die Wirkung eines internen »Ruherausens« ergibt, so kann durch entsprechend unterschiedliche Dimensionierung der Quantisierungen innerhalb der bandpaßbegrenzten PCM-Kanäle das Leerkanalrauschen der PCM-Strecke optimal an das Ruherausens des Gehörs angepaßt werden. Nimmt man ebenso an, daß sich die Mithörschwelle durch die Wirkung eines internen »Codierungsrauschens« ergibt, so kann durch entsprechend unterschiedliche Dimensionierung der Kompandierung innerhalb der bandpaßbegrenzten PCM-Kanäle das Kompandierungsrauschen der PCM-Strecke optimal an das Codierungsrauschen des Gehörs angepaßt werden.

Da die Steilheiten der Ruhehörschwellenkurven wesentlich kleiner sind als die Steilheiten der Mithörschwellenkurven, ist die erforderliche Breite der Bandpässe bestimmt durch die Verdeckung des Kompandierungsrauschens. Die Verdeckung des Leerkanalrauschens ist weniger kritisch. Optimal sind offenbar frequenzgruppenbreite Bandpässe, die das Kompandierungsrauschen im hohen Maße verdecken, und die deshalb hohe Kompandierungsgrade ermöglichen. Dies ist eine Frage des technischen Aufwandes.

Anpassung an die Amplitudenstatistik

Um den Aussteuerungsbereich einer PCM-Strecke auch bei hohen Frequenzen vollständig zu nutzen, benutzt man Pre/Deemphasen, die entsprechend der Amplitudenstatistik besonders im Bereich hoher Frequenzen wirksam sind. Die gebräuchlichen Pre-Deemphasen vermindern die Wahrnehmbarkeit des Quantisierungsrauschens, indem sie den Rauschpegel im Bereich hoher Frequenzen herabsetzen.

Langzeit-Spektrum

In Fig. 2 sind die Verhältnisse an zwei Beispielen dargestellt. Eingetragen sind die Ruhehörschwelle, dazu das Spektrum von weißem Rauschen (W.R.), das gerade noch unschwellig ist, sowie die resultierenden Spektren nach Bewertung durch die »50 μ s- bzw. J17-Pre/Deemphasis«. Die Pre-Deemphasen vermindern den Rauschpegel im Bereich der hohen Frequenzen und verschieben dadurch den Rauschpegel, der gerade noch verdeckt ist, um die Differenz ΔL , wobei ΔL der Ruhehörschwellenabstand ist. Für die 50 μ s-Pre/Deemphasis ist $\Delta L_{50} = -4$ dB, für die J17-Pre/Deemphasis ist $\Delta L_{J17} = -12$ dB, woraus sich Einsparungen von 21 kbit/s (50 μ s) bzw. 64 kbit/s (J17) ergeben.

Man erkennt anhand von Fig. 2, daß die Höhe der möglichen Einsparungen an Kanalkapazität in Breitband-PCM-Systemen allein durch das Maß der Vorverzerrung im 3 kHz-Bereich bestimmt ist. Allein der empfindliche Bereich des Gehörs in diesem Frequenzbereich ist entscheidend dafür, ob das Quantisierungsrauschen über- oder unschwellig ist.

Im Prinzip ergibt sich aber gerade im Bereich der hohen Frequenzen die größtmögliche Vorverzerrung. Besonders am Beispiel 50 μ s-Pre/Deemphasis wird deutlich, daß die Vorverzerrung bei Breitband-PCM nicht optimal zur Einsparung an Kanalkapazität genutzt wird. Im Vergleich zur frequenzunabhängigen Clippgrenze der PCM-Strecke ist rechnerisch eine Bitraten-Reduktion von 41 kbit/s möglich, wenn die Clippgrenze den Verlauf der 50 μ s-Deemphasis aufweist. Diese Reduktion ist durch Frequenzaufteilung technisch durch beispielsweise acht terzbreite Bandpässe im oberen Frequenzbereich erreichbar.

Offenbar kann die in Bandpaß-PCM-Systemen mögliche frequenzabhängige Quantisierung nicht nur zum Zwecke der Rauschanpassung vorteilhaft genutzt werden, sondern auch für die Anpassung des Kanals an die Amplitudenstatistik. Bandpaß-PCM-Systeme können die Amplitudenstatistik effektiver nutzen als Breitband-PCM-Systeme.

Für die Beantwortung der Frage, welcher Clippgrenzen-Frequenzverlauf tolerierbar ist, ist folgendes zu beachten:

Wie neuere amplitudenstatistische Untersuchungen zeigen, können bei moderner Musik insbesondere bei elektronischer Musik und impulshaften Klängen, die 50 μ s-Grenzkurve und sogar die 25 μ s-Grenzkurve zu kurzzeitigen Übersteuerungen führen, wenn ein Aussteuerungsmesser nach ARD-Pflichtenheft 3/6 benutzt wird. Bei Breitband-PCM-Systemen entstehen durch das Clipping kurzzeitige starke, nichtlineare Verzerrungen, die mit ausreichender Aussteuerungsreserve, also auf Kosten einer Verringerung des Rauschabstandes, vermieden werden. Aus diesem Grunde erscheint eine Herabsetzung der Clippgrenze entsprechend der J17-Kurve in Breitband-PCM-Systemen kaum sinnvoll. Indessen sind die Nachteile einer (weitgehenden) Pre/Deemphasis in Bandpaß-PCM-Systemen geringer: Infolge der Bandpaßbegrenzung werden die Klirrprodukte, die durch Clipping kurzzeitig auftreten, zum Teil unterdrückt und wesentlich besser verdeckt. Die Clipping-Geräusche werden – abhängig von der Dimensionierung der Bandpässe – mehr oder weniger stark gemildert. Besonders in schmalbandigen Bandpässen treten deutlich geringere Aussteuerungsprobleme auf, so daß im Bereich hoher Frequenzen die Clippgrenze wirkungsvoll herabgesetzt werden darf. Bereits oktav-breite Bandpässe unterdrücken praktisch alle harmonischen Verzerrungsprodukte, die durch Clipping auftreten.

Kurzzeit-Spektrum

Darüber hinaus besteht die Möglichkeit, das Prinzip »variable Preemphasis« anzuwenden. Eine adaptive Emphasis kann die Aussteuerungsprobleme weiter reduzieren, wenn Pre- und Deemphasis durch Übertragung der Steuergröße invers zueinander arbeiten. Der dafür erforderliche Informationsfluß ist gering.

Im Gegensatz zu Breitband-PCM-Systemen kann in Bandpaß-PCM-Systemen diese frequenzabhängige, kurzzeitige Anpassung sehr einfach erfolgen. Der momentan wirksame Frequenzverlauf der Preemphasis muß nicht

durch Kurzzeit-Spektralanalyse ermittelt und durch entsprechende variable Filter eingestellt werden, sondern es erfolgt eine einfache komplementäre Verstärkungsregelung⁶ in jedem Kanal der Bandpaß-PCM-Strecke. Dabei ist von Vorteil, daß die für die Hörschwellenanpassung optimierte Bandpaß-Aufteilung auch eine günstige Bandpaß-Aufteilung in Hinblick auf Amplitudenstatistik-Anpassung darstellt.

Eine derartige adaptive komplementäre Kanalverstärkung, die abhängig vom Pegel des Nutzsignals arbeitet, kann beispielsweise in der Weise ausgelegt sein, daß nur im Fall sehr hoher Nutzsignalpegel im Coder eine relativ hohe Kanalverstärkung zurückgeregelt wird (Limiter-Kennlinie). In diesem Fall ergibt sich die Rauschunterdrückung aus der Pegelanhebung im Coder und der komplementären Pegelabsenkung im Decoder; sie reduziert sich im Fall der Übersteuerung entsprechend der Limiter-Kennlinie.

Die adaptive komplementäre Kanalverstärkung kann aber auch in der Weise ausgelegt sein, daß im Fall niedriger und mittlerer Nutzsignalpegel im Coder die Kanalverstärkung hochgeregelt wird (Dynamikkompressor-Kennlinie). In diesem Fall ergibt sich durch die komplementäre Kanalverstärkung im Decoder für niedrige und mittlere Pegel ein Kompandergewinn entsprechend der Kompressor-Kennlinie.

Bei Anwendung von adaptiven komplementären Verstärkungen in den Kanälen der Bandpaß-PCM-Strecke zeigt sich ein weiterer Vorteil der Bandpaß-PCM gegenüber Breitband-PCM, weil sich die bekannten Kennlinien (Limiter, Kompressor, Gleitkomma) günstig an das Kurzzeit-Spektrum anpassen lassen. Die statischen und dynamischen Kennwerte der Kanalverstärkungen lassen sich dadurch auf einfachem Wege so optimieren, daß gegenüber Breitbandsystemen eine wesentlich höhere Rauschverminderung möglich ist.

Prinzipieller Aufbau der Bandpaß-PCM

In Fig. 3 ist ein prinzipieller Aufbau der Bandpaß-PCM dargestellt. Die Frequenzumsetzung und Bandpaßfilterung geschieht zur besseren Verständlichkeit analog. Schaltungstechnisch sind diverse digital arbeitende sogenannte »Transmultiplexer« realisiert, die Frequenzmultiplexsignale (FDM) umsetzen in Zeitmultiplexsignale (TDM) und umgekehrt. Man sieht leicht ein, daß ein breitbandiges Analogsignal der Breite B im Prinzip als Frequenzmultiplexsignal der Breite $B = \sum B_i$ aufgefaßt werden kann. Einige Verfahren werden beispielsweise in der Zeitschrift »The Bell System Technical Journal« Oktober 1976, Seiten 1069 bis 1085 und Mai-Juni 1977, Seiten 747 bis 770 sowie in der Dissertation »Dimensionierung von digitalen TDM-FDM-Transmultiplexern nach der Polyphasenmethode«, ETH Zürich 1979 diskutiert und beschrieben.

Anhand des prinzipiellen Aufbaus in Fig. 3 wird nachstehend erläutert, wie durch Wahl der Frequenzbandaufteilung (Anzahl Mittenfrequenz und Breite der Bandpässe) sowie durch Auslegung der Quantisierung die Verringerung der Übertragungs-Bitrate ohne Verschlechterung der Signalqualität erreicht werden kann.

Nach der Bandpaß-Aufteilung werden alle höher gelegenen Kanäle frequenzumgesetzt, so daß in der nachfolgenden PCM-Umsetzung die Abtastfrequenz $f_{st} \geq 2 B_i$ ausreicht. Der Netto-Informationsfluß I_{TDM} des Zeitmultiplex-Signals ist im Prinzip nicht größer als der Netto-Informationsfluß I_{FDM} des Frequenzmultiplex-Signals, er reduziert sich mit Verkleinerung der Rauschabstände in den einzelnen Kanälen, die nach Maßgabe der Hörschwellenkurven des Gehörs möglich sind.

Auf der Decoderseite wird das TDM-Signal invers zur Coderseite zurücktransponiert. Infolge der Bandpässe an den Kanalausgängen wird das Quantisierungsrauschen jedes Kanals bandpaßbegrenzt. Wenn die Quantisierung in allen Kanälen gleich ist, ergibt sich am Ausgang des Decoders das übliche (breitbandige) Rauschen einer Breitband-PCM-Strecke. Wird in einem der Kanäle die Auflösung beispielsweise verringert, so erhöht sich dementsprechend die Rauschleistung der Bandpaß-PCM-Strecke bei den Spektralanteilen, die im Durchlaßbereich dieses Kanals liegen.

Anpassung des Leerkanalrauschens an die Ruhehörschwelle

Fig. 4 zeigt die Ruhehörschwelle im Terzdiagramm. Die schraffierte Fläche stellt das ausnutzbare Spektrum dar. Zur Übertragung dieses Spektrums ist eine Kanalkapazität von etwa 60 kbit/s erforderlich (Grenzfrequenz 15 kHz). Der wesentliche Teil dieser Kapazität resultiert aus dem hochfrequenten Anteil des Spektrums. Das wird deutlich in der Darstellung in Fig. 5. Aufgetragen ist die mögliche Einsparung an Kanalkapazität je Terz, berechnet nach der Beziehung

$$C_{\text{Einsparung}} = \frac{1}{3} B_i a_{qi}' \quad (1)$$

B_i : Bandbreite des Bandpasses i mit Mittenfrequenz f_i
 a_{qi}' : Verringerung des Rauschabstandes (dB) im Bandpaß i

Der tieffrequente Anteil liefert keinen nennenswerten Beitrag (wegen der geringen Bandbreite B_i dieser Terzen). Allein mit den letzten drei Terzen 8, 10, 12,5 kHz einschließlich 1/4 der 16 kHz-Terz kann eine Einsparung von 48 kbit/s erzielt werden, das sind 80% der möglichen Gesamteinsparung. Dieser Betrag ergibt sich ebenso, wenn die letzte Oktave des gesamten Spektrums, nämlich 7,5 ... 15 kHz zugrundegelegt wird: es ist mit $B_i = 7,5$ kHz und $a_{qi}' = 19$ dB (vgl. Ruhehörschwelle) $C_{\text{Einsparung}} = 48$ kbit/s.

Als technisch sinnvolle Möglichkeit der Einsparung durch Anpassung des Leerkanalrauschens an die Ruhehörschwelle bietet sich also eine Halbierung des Frequenzbandes an. Diese Lösung nutzt 80% der möglichen Einsparung und erfordert für die Abtastung nur eine gemeinsame Abtastfrequenz von $f_s = 16$ kHz. Auf diesem Wege können damit 48 kbit/s ohne Qualitätsverlust eingespart werden.

Eine gemeinsame Abtastfrequenz für alle Kanäle der Bandpaß-PCM bietet technische Vorteile besonders bei der Wort- und Blocksynchronisation.

Anpassung des Kompandierungsrauschens an die Mithörschwelle

Nachstehend wird erläutert, weshalb allein durch Wahl der Breite B_f des Bandpasses / der Rauschabstand a_{qi} festgelegt wird, der erforderlich ist, um vollständige Verdeckung zu erreichen.

Die Verdeckungskurven sind, wenn sie statt über der Frequenz über der Barkskala aufgetragen werden, unabhängig von der Tonheit z . Oberhalb etwa 500 Hz ist $z \sim \lg f$, d. h. auch bei hohen Frequenzen können wir etwa die gleichen Kurvenverläufe annehmen wie bei mittleren Frequenzen, vgl. Fig. 1. Die Verdeckungswirkung eines Maskierers verhält sich oberhalb 500 Hz in 1. Näherung unabhängig von der Frequenzlage des Maskierers.

Weiterhin ist festzustellen, daß die sehr steilen unteren Flanken der Mithörschwellen darüber hinaus unabhängig sind vom Schalldruck des Maskierers (oberhalb 500 Hz), die Steilheit beträgt konstant etwa 80 ... 90 dB/Okt.

Um die Verdeckung eines bandpaßbegrenzten Kompandierungsrauschens abschätzen zu können, sei das worst-case-Beispiel entsprechend Bild 6 angenommen. Dargestellt ist die Mithörschwellenkurve für $f_M = 1$ kHz, Schalldruckpegel $L_M = 60$ dB. Beträgt die obere Grenzfrequenz des Bandpaßfilters $f_o = 1$ kHz, dann liegt die untere Grenzfrequenz des Kompandierungsrauschens bei $f_u = f_o - B_f$. In der Skizze ist B_f terzbreit angenommen, hier muß der Pegel des Kompandierungsrauschens etwa 25 dB unter dem Maskiererpegel liegen («kritischer Rauschabstand» für terzbreites Rauschen).

Rechnerisch ergibt sich der kritische Rauschabstand a_{qi} aus der Steilheit der unteren Flanke der Mithörschwellenkurve sowie aus der maximalen Höhe der Mithörschwellenkurve. Legt man die »Mithörschwellen eines Tones, verdeckt durch Schmalbandrauschen« zugrunde, so beträgt die Steilheit der unteren Flanke für $f_M = 1$ kHz 85 dB/Okt. und das Maximum $L_M - L_{Tmax} \approx 5$ dB (vgl. Bild 6). Damit ergibt sich für den kritischen Rauschabstand

$$\overline{a_{qi}} \approx 280 \lg \left(\frac{f_{oi}}{f_{ui}} \right) + 5 \text{ (dB)} \quad (2)$$

f_{oi}, f_{ui} : obere und untere Grenzfrequenz des Bandpasses /

Die Beziehung ist in erster Näherung unabhängig von der Frequenz und vom Schalldruck. Das trifft nicht mehr zu, wenn die Mithörschwellenkurven für maskierende Sinustöne zugrunde gelegt werden: Die Steilheit der unteren Flanke beträgt ca. 30 dB/Okt. für $L_M = 30$ dB und ca. 110 dB/Okt. für $L_M = 90$ dB; außerdem ist $L_M - L_{Tmax} \approx 11$ dB. Jedoch ergeben sich aus der oben angegebenen Beziehung erst bei Sinuspegeln $L_M > 60$ dB etwas zu kleinere Werte für den kritischen Rauschabstand a_{qi} .

Diese Darstellung basiert auf dem theoretisch ungünstigsten Fall. In der Praxis liegen die Verhältnisse aus zwei Gründen günstiger.

1. Die Übertragung einer einzelnen Linie genau an der oberen Grenzfrequenz f_{oi} des Bandpasses kommt praktisch nicht vor. Würde sich die Linie etwa in der Mitte des Durchlaßbereichs befinden, so würde der kritische Rauschabstand bereits halbiert werden.
2. Die vollkommen rauschfreie Übertragung einer einzelnen Linie und Wiedergabe mit Schalldruckpegeln > 60 dB ist selbst für elektronische Musik kaum relevant.

Für die worst-case-Abschätzung soll deshalb der kritische Rauschabstand angenommen werden mit

$$\overline{a_{qi}} = 200 \lg \left(\frac{f_{oi}}{f_{ui}} \right) + 5 \text{ (dB)} \quad (3)$$

Diesem Wert steht bei q_i Quantisierungsstufen ein Rauschabstand

$$\overline{a_{qi}} = 20 \lg (\sqrt{15} \cdot q_i) \text{ (dB)} \quad (4)$$

gegenüber. Mit Gleichung (3) und (4) läßt sich die für die Verdeckung des Kompandierungsrauschens erforderliche Bitrate direkt bestimmen.

$$ld\overline{q_i} = 33,2 \lg \left(\frac{f_{oi}}{f_{ui}} \right) + 0,52 \left[\frac{\text{bit}}{\text{sample}} \right] \quad (5)$$

Für die Abschätzung wird angenommen:

$$ld\overline{q_i} = 32 \lg \left(\frac{f_{oi}}{f_{ui}} \right) + 1 \left[\frac{\text{bit}}{\text{sample}} \right] \quad (6)$$

Daraus ergibt sich beispielsweise folgende Wertetabelle:

Tabelle 1:

| | | | | |
|---|--|--|---|----|
| Bandbreite B_f : | 1 Terz | 2 Terzen | 3 Terzen | 5 |
| krit. Rauschabstand \overline{a}_{qr} : | 25 dB | 45 dB | 65 dB | |
| Für Verdeckung erf. $ld\overline{q}$ | 4,1 $\frac{\text{bit}}{\text{sample}}$ | 7,4 $\frac{\text{bit}}{\text{sample}}$ | 10,6 $\frac{\text{bit}}{\text{sample}}$ | 10 |

Man erkennt die günstige Wirkung eines schmalen Bandpasses, damit aber auch die ungünstigen Verhältnisse bei Breitband-Kompandierung. Wenn die Bandbreiten größer als 4 Terzen sind, verdeckt die Mithörschwelle nicht mehr das Kompandierungsrauschen, sondern sie reduziert lediglich die Störwirkung.

An dieser Stelle sei bereits auf einen günstig wirkenden Sachverhalt hingewiesen. Man kann in der Praxis davon ausgehen, daß die spektralen Anteile im Bereich hoher Frequenzen, beispielsweise ab etwa 7,5 kHz, nur als Oberton- oder Geräuschanteil eines Gesamtklanges auftreten. Selbst elektronische Musikinstrumente erzeugen kaum Klänge mit nur hochfrequenten Anteilen. Deshalb wird im Bereich hoher Frequenzen die Verdeckung des Kompandierungsrauschens im Bandpaß / vorwiegend durch Nutzsignalanteile im Bandpaß / - 1 erfolgen. Da die oberen Flanken der Mithörschwellen wesentlich flacher verlaufen, ist die für Verdeckung erforderliche Bitrate im Bereich hoher Frequenzen praktisch kleiner als nach Abschätzung entsprechend Gleichung (6).

Im Hinblick auf eine günstige Anpassung des Bandpaß-PCM-Kanals an den Ohrkanal soll festgehalten werden:

- Die für Verdeckung erforderliche Bitrate ist praktisch nur von der Bandpaßbreite abhängig, dagegen praktisch nicht von der Bandpaß-Mittenfrequenz und der Abhörlautstärke.
- Ein Bandpaß-PCM-System, das den kritischen Rauschabstand einhält, erzeugt auch bei hohen Abhörpegeln kein hörbares Kompandierungsrauschen.
- Ein Bandpaß-PCM-System, das den kritischen Rauschabstand unterschreitet, kann ein hörbares Kompandierungsrauschen erzeugen, wenn das Kompandierungsrauschen über der Ruhehörschwelle liegt.

Liegt der kritische Rauschpegel unterhalb der Ruhehörschwelle, so verdeckt die Ruhehörschwelle. In diesem Fall ist entweder der Kompandierungsgrad unnötig hoch oder der Bandpaß unnötig schmal. Für die Dimensionierung der Bandpaß-PCM ist also im Prinzip die maximale Abhörlautstärke von Bedeutung.

Anpassung des Bandpaß-PCM-Kanals an den Ohrkanal

Für die Bestimmung einer günstigen Anpassung des Bandpaß-PCM-Kanals an den Ohrkanal wird das »Signal-Rauschverhältnis des Ohrkanals« in Abhängigkeit vom Frequenzbereich festgestellt und in den bekannten S/N-Diagrammen dargestellt. Dazu ist die Abhörlautstärke festzulegen. Es wird angenommen, daß bei Vollaussteuerung der PCM-Strecke der Schalldruckpegel 90 dB beträgt. Der Nominalpegel $S_0 = 0$ dB der PCM-Strecke ergibt also einen Nominalschalldruckpegel von $L_{S0} = 90$ dB. Weiter wird angenommen, daß der headroom 10 dB groß ist.

Mit Blick auf den Frequenzverlauf der Ruhehörschwelle wird die Anpassung für drei »typische« Mittenfrequenzen ermittelt: $f_{m1} = 0,8$ kHz, $f_{m2} = 3,2$ kHz, $f_{m3} = 10$ kHz; Fig. 7 zeigt die angenommenen Verhältnisse in der Hörfläche. Eingetragen sind auch die Mithörschwellen für Vollaussteuerung sowie die kritischen Rauschabstände für eine, zwei und drei Terzen (1T, 2T, 3T), vgl. Tabelle 1.

Man sieht, daß der kritische Rauschabstand bei geringerem Schalldruckpegel, also bei geringerer Aussteuerung der PCM-Strecke, unterhalb der Ruhehörschwelle liegen kann. Beispielsweise ist der erforderliche Rauschabstand eines 10-kHz-Tones nicht mehr gleich dem kritischen Rauschabstand $(\overline{a}_{q10\text{ kHz}})_{3T} = 65$ dB, wenn der Pegel des 10-kHz-Tones kleiner wird als 10 dB unter Vollaussteuerung: hier wirkt schon die Ruhehörschwelle »verdeckend«. Wird der Pegel weiter reduziert, so darf der Rauschabstand im 3-Terz-breiten Bandpaß proportional mit dem Signalpegel S fallen.

Dieser Verlauf des Rauschabstandes des Gehörs ist der Verlauf des erforderlichen Rauschabstandes des PCM-Kanals. Er kann, wie üblich, in einem S/N-Diagramm dargestellt werden, und aus der Beziehung Nominalschalldruckpegel/Nominalpegel ergeben sich direkt die Anforderungen an den bandpaßbegrenzten Kanal bezüglich Leerkanalrauschen und Kompandierungsrauschen.

Im Diagramm Fig. 8 sind die 45°-Geraden eingetragen, die sich aus Nominalschalldruckpegel und Ruhehörschwelle ergeben, sie sind frequenzabhängig. Weiterhin sind die kritischen Rauschabstände für die Bandpaßbreiten 1T, 2T und 3T (1, 2 und 3 Terzen) eingetragen, die sich aus der Mithörschwelle ergeben; sie sind frequenzunabhängig. Die Kurven beschreiben exakt den »Rauschabstand des Gehörs« $(S/N)_{Oh}$ (d. h., den Abstand eines Tones von der Hörschwelle im Bereich eines Bandpasses): Im Bereich kleiner Schalldruckpegel L_S wächst der Abstand zur (Ruhe)-Hörschwelle des bandpaßbegrenzten Rauschens proportional mit L_S an. Im Bereich hoher Schalldruckpegel L_S (abhängig von der Bandpaßbreite) bleibt der Abstand des bandpaßbegrenzten Rauschens von der (Mit)-Hörschwelle konstant.

Ein derartiger (S/N)-Verlauf kennzeichnet bekanntlich auch PCM-Strecken mit nichtlinearer, insbesondere logarithmischer Quantisierung (z. B. 13-Segment-Kompanderkennlinie): Im Bereich kleiner Signalpegel S

wächst der Abstand $(S/N)_{PCM}$ proportional mit S an; im Bereich hoher Signalpegel bleibt $(S/N)_{PCM}$ in erster Näherung konstant.

Die optimale Quantisierung für den Bandpaß i der Bandpaß-PCM-Strecke läßt sich demnach unmittelbar den $(S/N)_{Ohf}$ -Verläufen entnehmen. Fig. 8 macht zunächst wichtige Zusammenhänge deutlich:

1. Die Lage der waagerechten Geraden kennzeichnet den maximal erforderlichen Rauschabstand $(S/N)_{PCM}$. Er hängt allein von der Bandbreite des Bandpaß-Kanals ab:
Es gilt entsprechend Gleichung (3):

$$(S/N)_{PCM \max} = \overline{a_{qi}} = 200 \lg \left(\frac{f_{oi}}{f_{ui}} \right) + 5 \text{ (dB)} \quad (7)$$

2. Der maximal erforderliche Rauschabstand, und damit die erforderliche Anzahl der Bits je sample, ist unabhängig vom Nominalschalldruckpegel L_{SO} .
 3. Die Lage der 45°-Geraden kennzeichnet den Leerkanalrauschpegel, der an der Ruhehörschwelle liegt. Der Leerkanalrauschpegel verhält sich zur Höhe des Nominalschalldruckpegels L_{SO} umgekehrt proportional.
 4. Der Knickpunkt der Geraden kennzeichnet den Kompandergewinn bzw. den Kompandierungsgrad. Er hängt ab von der Bandbreite des Kanals i (1T, 2T oder 3T) und vom Nominalschalldruckpegel L_{SO} .
- Der Informationsfluß eines Bandpaß-PCM-Signals ist demnach ausschließlich durch die Breite und Anzahl der Bandpaß-Kanäle bestimmt. Er ist für $B_i \leq 4T$ unabhängig vom Nominalschalldruckpegel.
- Die Höhe des Nominalschalldruckpegels ist allein begrenzt durch die Wahrnehmbarkeit des Leerkanalrauschens im Kanal. Das Leerkanalrauschen vermindert sich proportional der Erhöhung des Kompandierungsgrades. Durch Dimensionierung der Kompandierkennlinie wird also die maximale Abhörlautstärke festgelegt, die gerade noch Rauschfreiheit gewährleistet.

Beispiel 1 (vgl. Bild 8)

Gewählt: $L_{so} = 90 \text{ dB}$, $B_i = 3T$

Dann ist: $(S/N)_{PCM \max} = 65 \text{ dB}$, entsprechend $10,6 \frac{\text{bit}}{\text{sample}}$.

Daraus resultieren Kompandergewinne K_i sowie Kompandierungsgrade G_i :

$f_{m1} = 0,8 \text{ kHz} : K_1 = 31 \text{ dB} \quad G_1 = 5,2 \text{ bit}$
 $f_{m2} = 3,2 \text{ kHz} : K_2 = 39 \text{ dB} \quad G_2 = 6,5 \text{ bit}$
 $f_{m3} = 10 \text{ kHz} : K_3 = 20 \text{ dB} \quad G_3 = 3,3 \text{ bit}$

Damit liegen die optimalen Quantisierungen für die 3 oktavbreiten Kanäle der Bandpaß-PCM-Strecke fest:

$f_{m1} = 0,8 \text{ kHz} : 15,7 \text{ auf } 10,6 \text{ bit/sample}$
 $f_{m2} = 3,2 \text{ kHz} : 17,0 \text{ auf } 10,6 \text{ bit/sample}$
 $f_{m3} = 10 \text{ kHz} : 13,8 \text{ auf } 10,6 \text{ bit/sample}$

Diese Bandpaß-Kanäle sind bis zu Spitzenschalldruckpegeln von $L_s = 100 \text{ dB}$ rauschfrei.

Würde man eine entsprechende Oktavband-PCM-Strecke aufbauen (8 Oktav-Bandpässe), so würde der resultierende Informationsfluß

$$I = \sum f_{si} \cdot 10,6 \text{ bit} = 340 \text{ kbit/s}$$

betragen. Gegenüber einer 16-bit-linear-PCM-Strecke (deren Systemdynamik 6 dB kleiner ist) bedeutet das eine Einsparung von 172 kbit/s (35%). Wollte man den für 16-bit-linear erforderlichen Informationsfluß halbieren (ca. 250 kbit/s), so müßte die Oktavbreite der Bandpässe um den Faktor 0,74 reduziert werden. Für die Erhaltung der Systemdynamik wären dann die Kompandierungsgrade um den Faktor 1,34 zu erhöhen.

Doch der spezifische Gewinn je bit Einsparung ist nur bei hohen Frequenzen groß. Deshalb ist es günstiger, im oberen Frequenzbereich wenige schmale Bandpässe vorzusehen, als im gesamten Frequenzbereich breitere Bandpässe. Bei gleichem Gewinn ist der technische Aufwand dann geringer. Erforderlich ist allerdings eine höhere Kompandierung.

Beispiel 2

Gewählt: $L_{so} = 90 \text{ dB}$
 $B_i = 1,5T$ (7,5 ... 10,7 und 10,7 ... 15 kHz)

Dann ist: $(S/N)_{PCM \text{ max}} = 35 \text{ dB}$, entsprechend $5,7 \frac{\text{bit}}{\text{sample}}$.

Hieraus ergibt sich ein Informationsfluß von 348 kbit/s; die Einsparung gegenüber 16-bit-linear beträgt 164 kbit/s. Dies wird erreicht durch Einteilung des Frequenzbandes in die drei Bänder 0,04 ... 7,5/7,5 ... 10,7/10,7 ... 15 kHz. Das untere Band arbeitet 16-bit-linear. Gegenüber Beispiel 1 erreicht man also mit drei Bändern praktisch den gleichen Informationsfluß. Diese und weitere Dimensionierungsbeispiele sind in Tabelle 2 zusammengestellt:

Tabelle 2

| Beispiel | i | f_{ul} kHz | f_{ol} kHz | bit sample | I_i kbit/s | $\sum I_i$ kbit/s | |
|----------|-----|-----------------|-----------------|---------------|-----------------|----------------------|----|
| 1 | 1 | 0,04 | 0,12 | 15,9 | 2,6 | 339,7 | 15 |
| | 2 | 0,12 | 0,25 | 10,6 | 2,7 | | |
| | 3 | 0,2 | 0,5 | 10,6 | 5,3 | | |
| | 4 | 0,5 | 0,9 | 10,6 | 10,6 | | |
| | 5 | 0,9 | 1,9 | 10,6 | 21,2 | | 20 |
| | 6 | 1,9 | 3,8 | 10,6 | 42,5 | | |
| | 7 | 3,8 | 7,5 | 10,6 | 84,9 | | |
| | 8 | 7,5 | 15,0 | 10,6 | 169,9 | | |
| 2 | 1 | 0,04 | 7,5 | 16,0 | 255,6 | 348,3 | 25 |
| | 2 | 7,5 | 10,7 | 5,9 | 40,5 | | |
| | 3 | 10,7 | 15,0 | 5,7 | 52,2 | | |
| 3 | 1 | 0,04 | 5,4 | 16,0 | 184,0 | 301,6 | 30 |
| | 2 | 5,4 | 7,5 | 5,6 | 24,9 | | |
| | 3 | 7,5 | 10,7 | 5,9 | 40,5 | | |
| | 4 | 10,7 | 15,0 | 5,7 | 52,2 | | |
| 4 | 1 | 0,04 | 5,0 | 16,0 | 170,4 | 354,2 | |
| | 2 | 5,0 | 10,0 | 10,6 | 113,2 | | |
| | 3 | 10,0 | 15,0 | 6,6 | 70,6 | | |
| 5 | 1 | 0,04 | 3,8 | 16,0 | 129,5 | 384,3 | 35 |
| | 2 | 3,8 | 7,5 | 10,6 | 84,9 | | |
| | 3 | 7,5 | 15,0 | 10,6 | 169,9 | | |
| 6 | 1 | 0,04 | 6,0 | 16,0 | 204,5 | 350,4 | 40 |
| | 2 | 6,0 | 9,0 | 6,6 | 42,4 | | |
| | 3 | 9,0 | 15,0 | 8,1 | 103,5 | | |
| 7 | 1 | 0,04 | 6,2 | 16,0 | 211,3 | 286,8 | 45 |
| | 2 | 6,2 | 7,7 | 4,1 | 12,8 | | |
| | 3 | 7,7 | 9,6 | 4,1 | 15,5 | | |
| | 4 | 9,6 | 12,0 | 4,1 | 21,0 | | |
| | 5 | 12,0 | 15,0 | 4,1 | 26,2 | | |

Die resultierenden Informationsflüsse $\sum I_i$ sind allein durch Anpassung der PCM-Strecke an die Mithörschwelle erreichbar. Man erkennt, daß bereits die Aufteilung in drei Frequenzbänder gegenüber 16-bit-linear eine Einsparung von etwa 150 kbit/s bewirken kann (Beispiele 2, 4, 6), und daß die Wahl der Grenzfrequenzen der drei Bandpässe nicht allein nach Maßgabe der Mithörschwelle festgelegt ist.

Anpassung des Bandpaß-PCM-Kanals an die Amplitudenstatistik

Deshalb ergibt sich im Hinblick auf die Anpassung an die Amplitudenstatistik ein Dimensionierungsspielraum. Die damit erreichbare zusätzliche Einsparung wird anhand der statischen J17-Pre/Deemphasis-Kurve erläutert; sie liegt entsprechend höher bei Einsatz einer adaptiven Anpassung an das Kurzzeit-Spektrum durch adaptive, komplementäre Verstärkung.

Fig. 9 zeigt ein Ausführungsbeispiel für die statische Anpassung an das Langzeit-Spektrum. Dargestellt ist die J17-Kurve (gestrichelte Linie), im Vergleich dazu die Ruhehörschwelle. Im Frequenzbereich 0,5 ... 6 kHz (Bandpaß $i = 1$) ist der Frequenzverlauf der Pre/Deemphasis genau der Ruhehörschwelle angepaßt («Ruhehörschwellen-Emphasis»). Im Bereich 6 ... 9 kHz (Bandpaß $i = 2$) erfolgt eine komplementäre Verstärkung von 12 dB, im Bereich 9 ... 15 kHz (Bandpaß $i = 3$) von 18 dB. Insgesamt wird der Aussteuerungsbereich dieser Bandpaß-PCM-Strecke im Frequenzbereich 1 ... 9 kHz um ca. 6 dB geringer eingeschränkt als im Fall der J17-Pre/Deemphasis bei Breitband-PCM-Strecken.

Dennoch ist die resultierende Informationsfluß-Einsparung noch um 12,8 kbit/s größer:

- Bandpaß 1: Der Ruhehörschwellen-Abstand beträgt 12 dB. Mit 2 bit/sample und einer Abtastfrequenz $f_s = 12,8$ kHz reduziert sich die Übertragungsrate um 25,6 kbit/s.
- Bandpaß 2: Die Dämpfung des Rauschpegels beträgt 12 dB. Mit 2 bit/sample und $f_s = 6,4$ kHz reduziert sich die Übertragungsrate um 12,8 kbit/s.
- Bandpaß 3: Die Dämpfung des Rauschpegels beträgt 18 dB. Mit 3 bit/sample und $f_s = 12,8$ kHz reduziert sich die Übertragungsrate um 38,4 kbit/s.

Dies sind insgesamt 76,8 kbit/s. Damit beträgt bei dieser Bandpaß-PCM-Strecke (Beispiel 6 in Tabelle 2) der Netto-Informationsfluß 274 kbit/s. Im Vergleich zu einer entsprechenden Breitband-PCM-Strecke (Quantisierung 16 auf 14, J17-Pre/Deemphasis) die einen Netto-Informationsfluß von 447 kbit/s übertragen muß, lassen sich durch die Aufteilung in nur drei Bandpässe bereits 173 kbit/s einsparen. Dabei ist im Vergleich zur J17-Pre/Deemphasis die Einschränkung des Aussteuerungsbereiches im Frequenzbereich 1...9 kHz um ca. 6 dB geringer, und die Tonqualität der beiden PCM-Systeme ist gleichwertig. Darüber hinaus bewirken die Bandpaß-Begrenzungen, daß die Verzerrungsprodukte, die durch Clipping bei Übersteuerung auftreten, im Frequenzbereich 6...15 kHz praktisch vollkommen unterdrückt werden. Kurzzeitige Übersteuerungen, die durch die Einschränkung des Aussteuerungsbereiches entsprechend Fig. 9 auftreten können, stören aus diesem Grunde und infolge der Verdeckung praktisch nicht. Im Gegensatz zur Breitband-PCM-Strecke macht deshalb die Nutzung der Langzeit-Amplitudenstatistik keine erhöhte Aussteuerungsreserve erforderlich.

Fig. 10 zeigt ein Ausführungsbeispiel für eine zusätzliche dynamische Anpassung an das Kurzzeit-Spektrum. Dargestellt ist eine oktavbreite Frequenzband-Aufteilung entsprechend Beispiel 1 in Tabelle 2. Der zu übertragende Informationsfluß dieser Bandpaß-PCM-Strecke beträgt ohne Nutzung der Amplituden-Statistik 340 kbit/s. Allein durch Einfügung von statischen komplementären Vorverstärkungen in den Bandpässen $l = 7$ und $l = 8$ von 6 bzw. 18 dB (schraffierte Bereiche und Clippgrenzenverlauf in Fig. 10) würde sich eine Einsparung von 8,0 kbit/s + 48,0 kbit/s = 56,0 kbit/s ergeben (Nutzung der Langzeit-Amplitudenstatistik). Sieht man darüber hinaus eine adaptive komplementäre Vorverstärkung mit 12 dB-Limiter-Kannlinie in jedem der Bandpässe $l = 5$ bis $l = 8$ vor (punktiierte Bereiche in Fig. 10), so reduziert sich der Informationsfluß auf 223,4 kbit/s:

Tabelle 3

| Beispiel | l | f_{ul} kHz | f_{ol} kHz | bit sample | I_l kbit/s | $\sum I_l$ kbit/s |
|----------|-----|-----------------|-----------------|---------------|-----------------|----------------------|
| 1a | 1 | 0,04 | 0,12 | 15,9 | 2,6 | 223,4 |
| | 2 | 0,12 | 0,25 | 10,6 | 2,7 | |
| | 3 | 0,25 | 0,5 | 10,6 | 5,3 | |
| | 4 | 0,5 | 0,9 | 10,6 | 10,6 | |
| | 5 | 0,9 | 1,9 | 8,6 | 17,2 | |
| | 6 | 1,9 | 3,8 | 8,6 | 34,5 | |
| | 7 | 3,8 | 7,5 | 7,6 | 60,9 | |
| | 8 | 7,5 | 15,0 | 5,6 | 89,6 | |

Diese Bandpaß-PCM-Strecke hat folgende Eigenschaften:

- Übersteuerungen (Überschreitung der Clippgrenze) im Frequenzbereich 0,1...15 kHz führen nicht zu störenden Geräuschen durch nichtlineare Verzerrungsprodukte, sondern zu weit geringer störenden kurzzeitigen Klangverfärbungen durch lineare Verzerrungen. Übersteuerungen sind unkritischer als bei analogen Übertragungsstrecken.
- Die Absenkung der Clippgrenze um 6 dB im Frequenzbereich 3,8...7,5 kHz und um 18 dB im Bereich 7,5...15 kHz (vgl. Fig. 10) vermindert deshalb auch bei kritischem Programm nicht die Aussteuerbarkeit der Strecke. Die Aussteuerungsreserve läßt sich sogar problemlos reduzieren.
- Kompandierungsrauschen ist auch bei größten Abhörpegeln nicht wahrnehmbar. Allein für reine Töne mit $f \geq 1,9$ kHz, deren Pegel größer sind als -12 dB (bezogen auf die Clippgrenze) ist theoretisch eine Überschreitung der Mithörschwelle durch das Kompandierungsrauschen möglich.
- Wenn durch entsprechende digitale Kompandierungen in allen acht Bandpässen die in Tabelle 3 vorgesehenen Bitraten (bit/sample) für die Übertragung eingehalten werden, so ist das Kompandierungsrauschen in der gesamten Bandbreite 0,04...15 kHz – unabhängig vom Kompandierungsgrad – unerschwellig. Deshalb kann ohne hörbare Verluste eine beliebig hohe Dynamik übertragen werden. Die maximale Dynamik ist somit nur durch die Systemdynamik der A/D-D/A-Wandler begrenzt.
- Infolge der Bandpaß-Begrenzungen und der komplementären Verstärkungen treten Störungen durch Bitfehler bei der Übertragung wesentlich seltener und schwächer auf als entsprechende Knackstörungen oder Störgeräusche bei Breitband-PCM-Übertragung.

Insgesamt ergeben sich für dieses Ausführungsbeispiel der erfindungsgemäßen Bandpaß-PCM-Strecke gegenüber einer entsprechenden Breitband-PCM-Strecke folgende Vorteile:

- Die Übertragungs-Bitrate für das Bandpaß-PCM-codierte AF-Signal beträgt etwa 230 kbit/s.
- Die maximale Dynamik ist nur durch die Auflösung der A/D-D/A-Wandler begrenzt.
- Das Kompondierungsrauschen ist unabhängig vom Abhörpegel nicht wahrnehmbar.
- Die Strecke ist unempfindlicher gegenüber Übersteuerungen als analoge Strecken.
- Die Strecke ist unempfindlich gegenüber Bitfehlern.

Hierzu 10 Blatt Zeichnungen

5

10

15

20

25

30

35

40

45

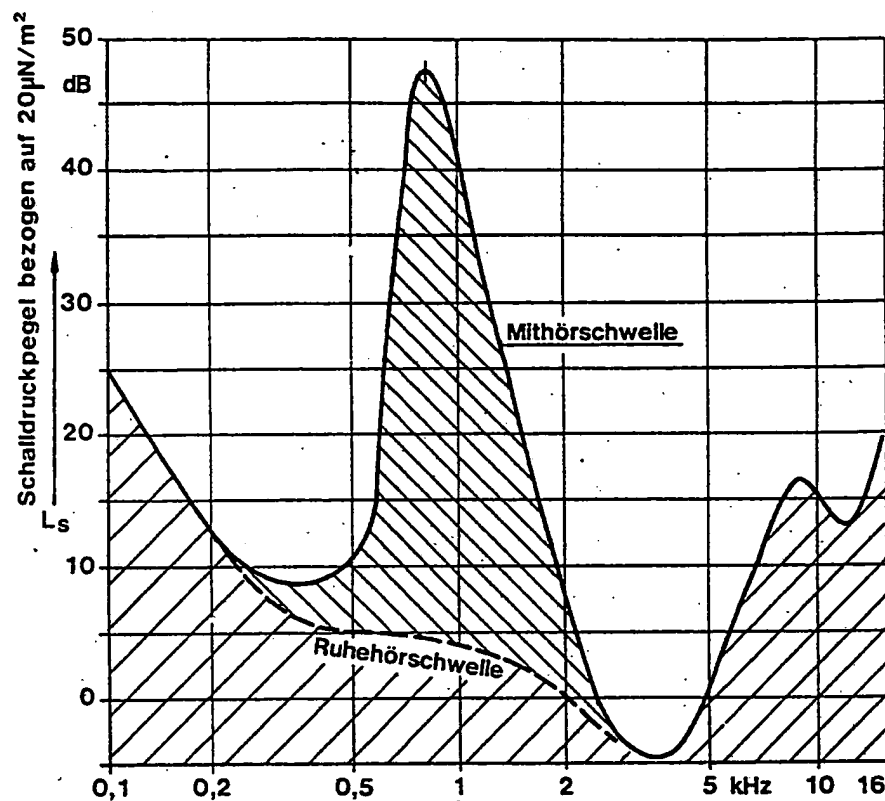
50

55

60

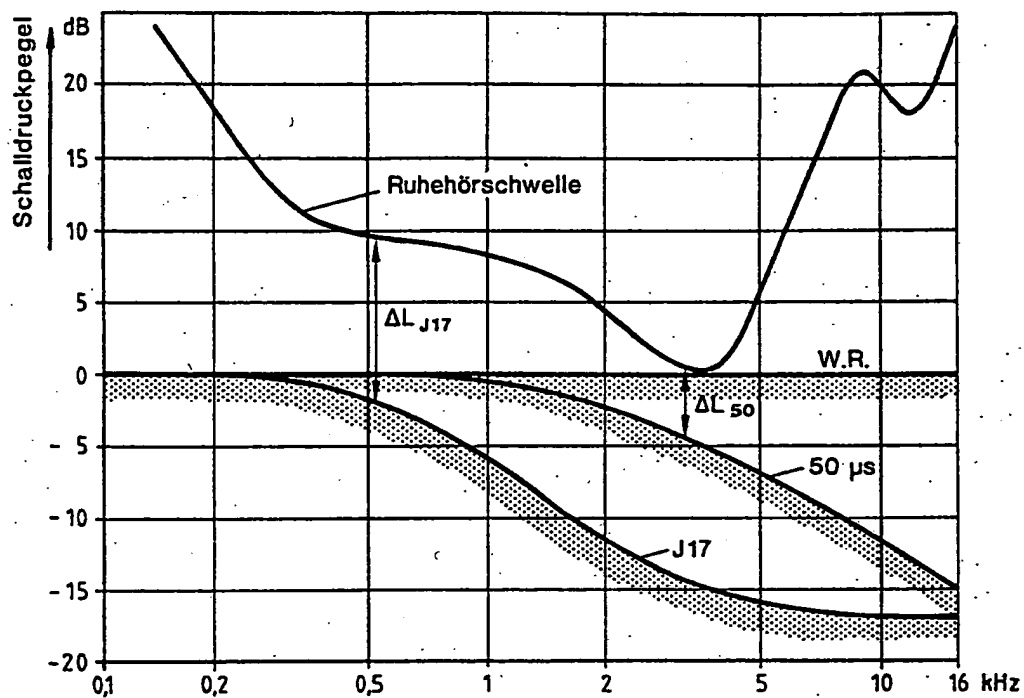
65

- Leerseite -



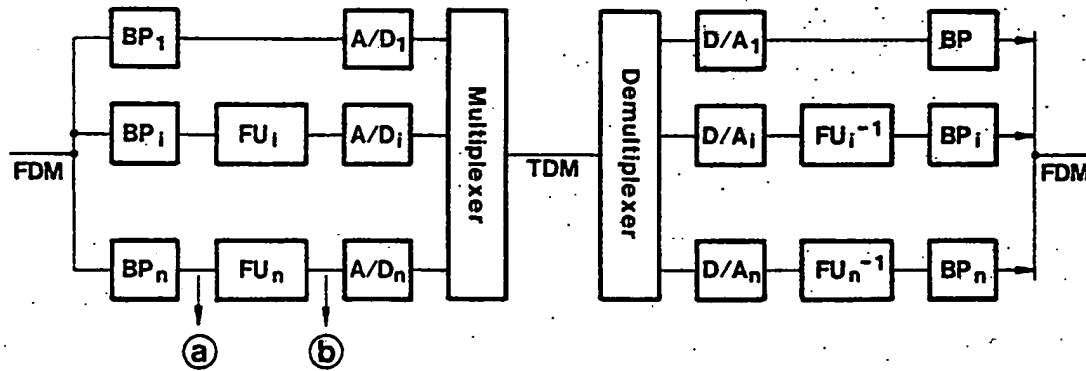
Hörschwellenkurven. Mithörschwelle für Schmalbandrauschen für 800Hz als Maskierer, 50dB Schalldruck. Der Schall im schraffierten Bereich ist unhörbar.

Fig. 1

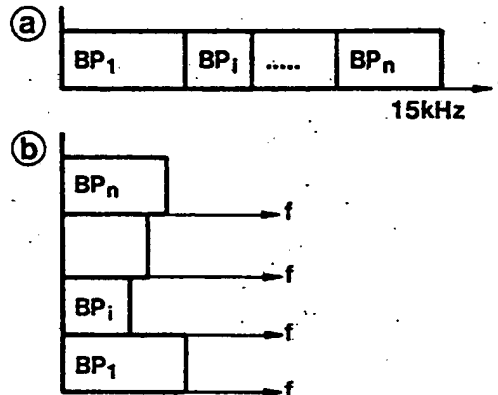


Verschiebung des Rauschpegels, der gerade noch verdeckt ist,
durch Frequenzbewertung entsprechend der 50µs- bzw. J17 - Pre/Deemphasis

Fig. 2

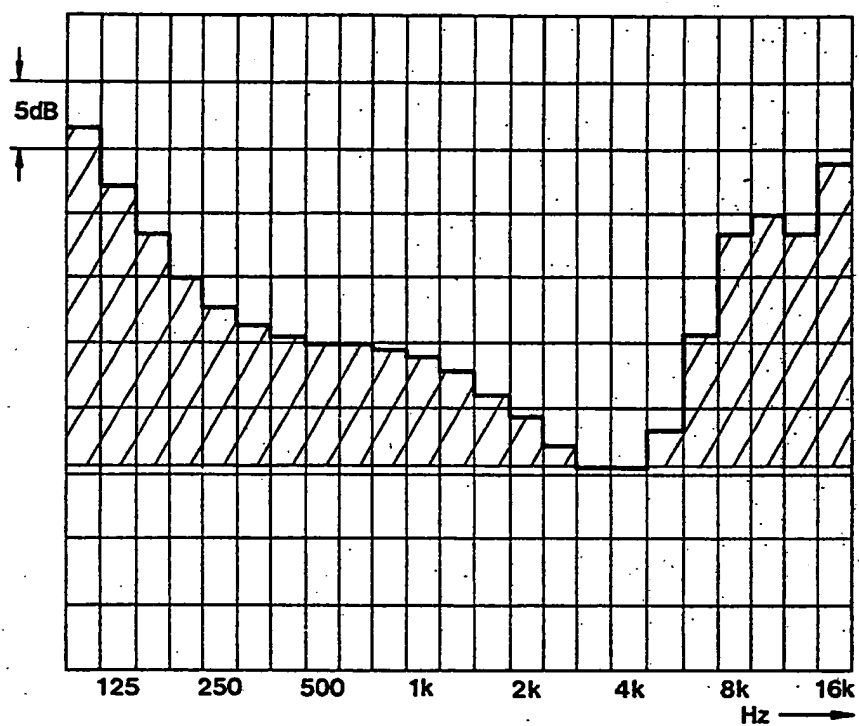


- 1...i...n Kanalnummern
 BP_i Bandpaß der Breite B_i
 FU_i Frequenzumsetzung
 A/D_i } PCM-Codierung/Decodierung,
 D/A_i } $f_s \approx 2B_i$
 beliebige Quantisierung



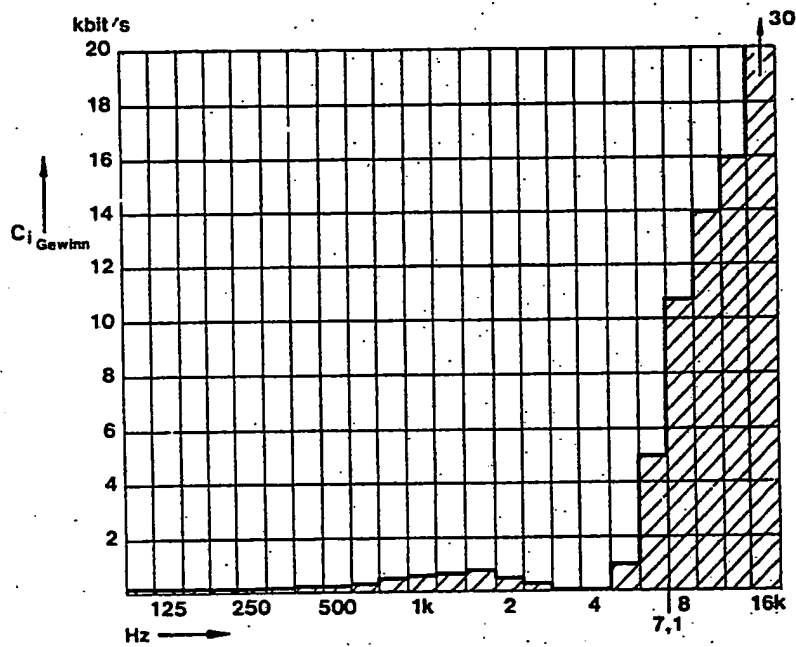
Prinzip einer Bandpaß-PCM-Strecke

Fig. 3



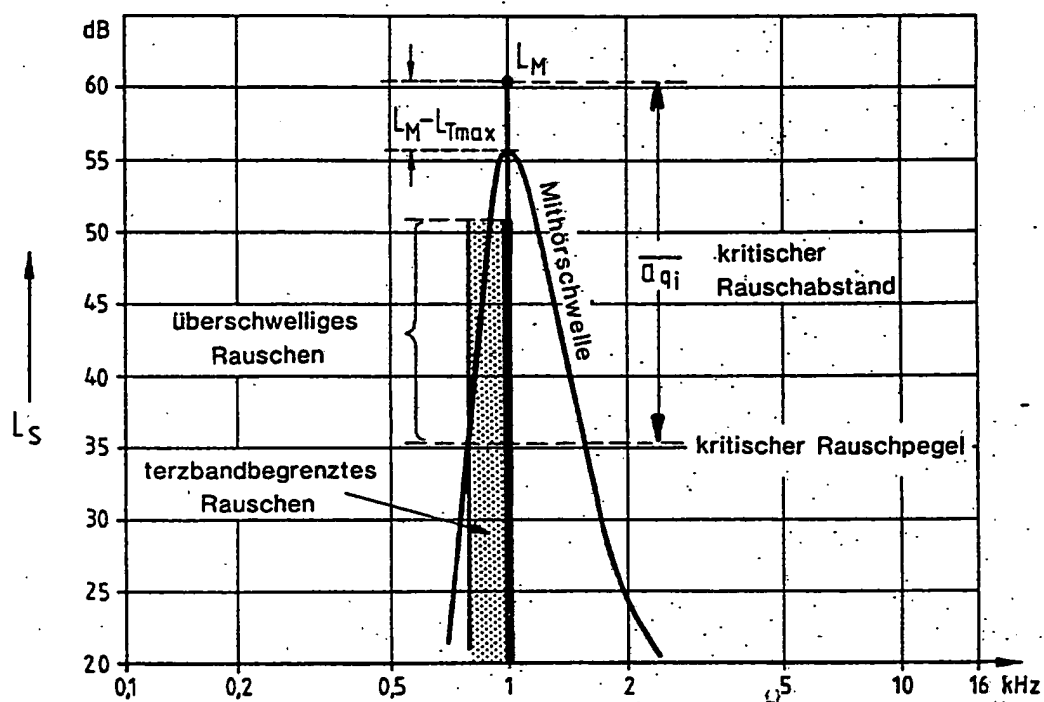
Ruhehörschwelle im Terzdiagramm. Für die Übertragung des Spektrums im schraffierten Bereich sind 60kbit/s erforderlich.

Fig. 4



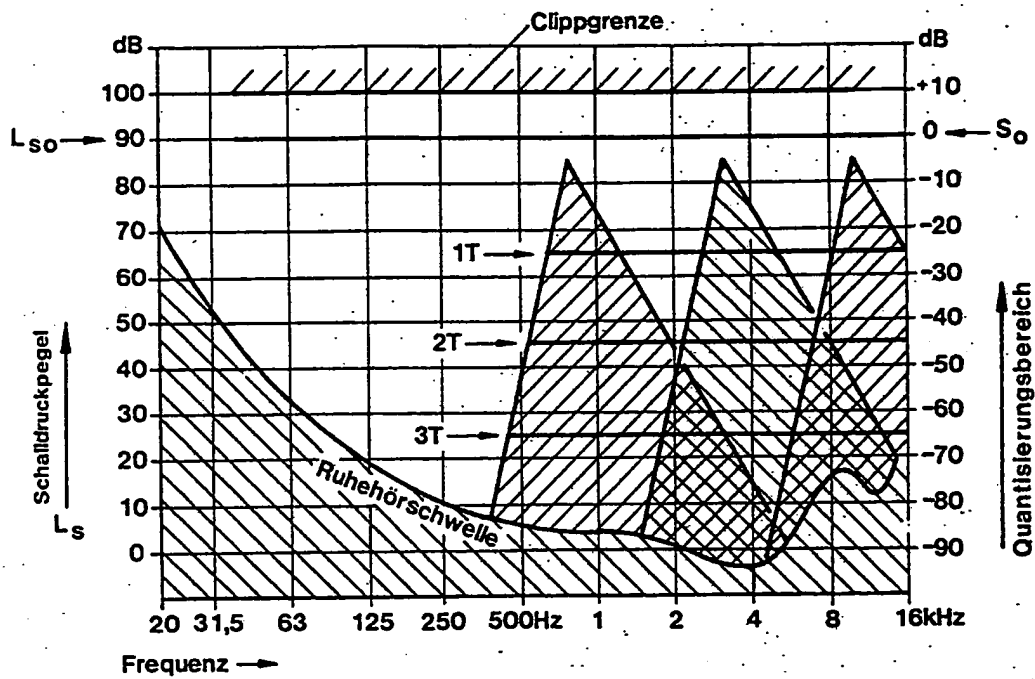
Einsparung an Kanalkapazität je Terz bei Anpassung des Leerkanal-ausgangs an die Ruhehörschwelle

Fig. 5



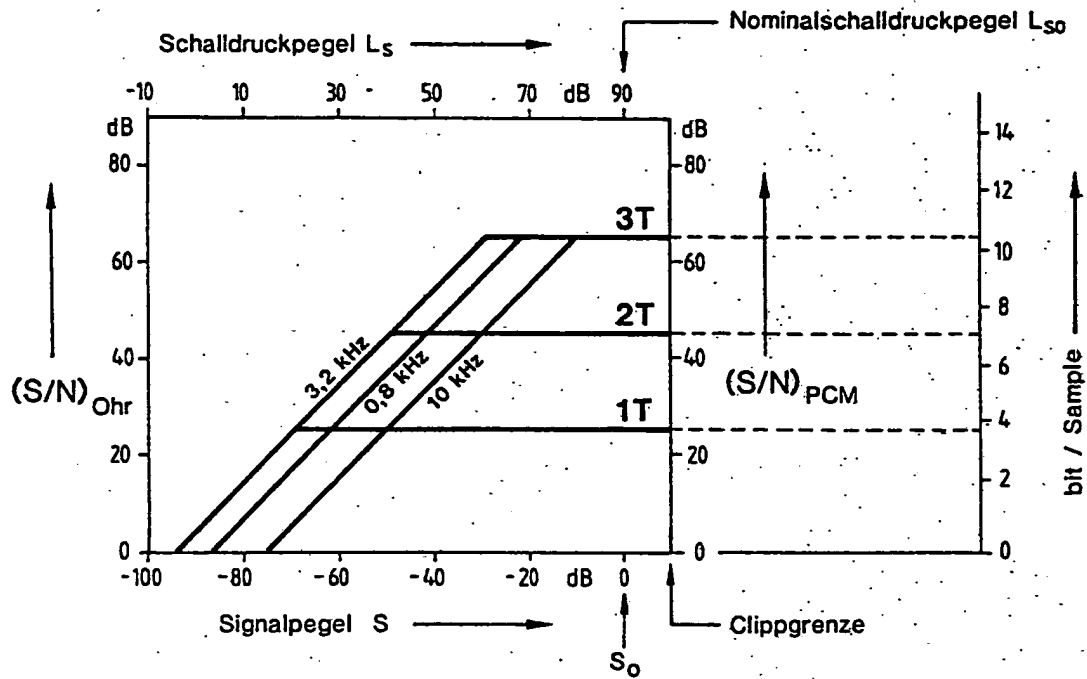
Skizze zur Abschätzung der Verdeckung
 von terzbandbegrenztem Kompandierungsrauschen

Fig. 6



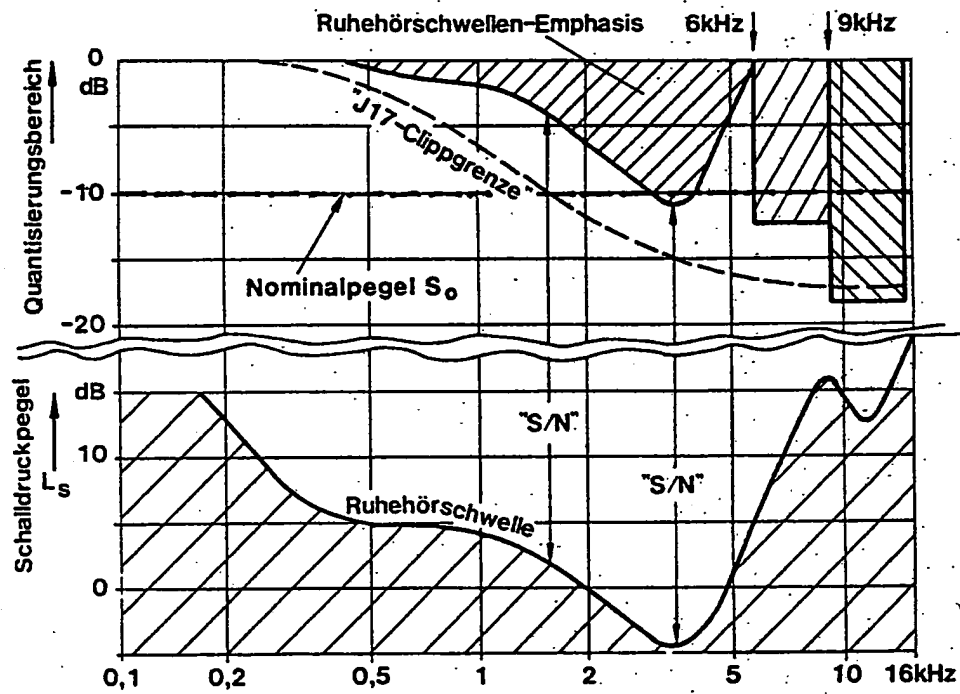
Ruherauschen und Codierungsrauschen des Gehörs in Bezug zum Leerkanalrauschen und Kompondierungsrauschen eines 16-bit-PCM-Systems

Fig. 7



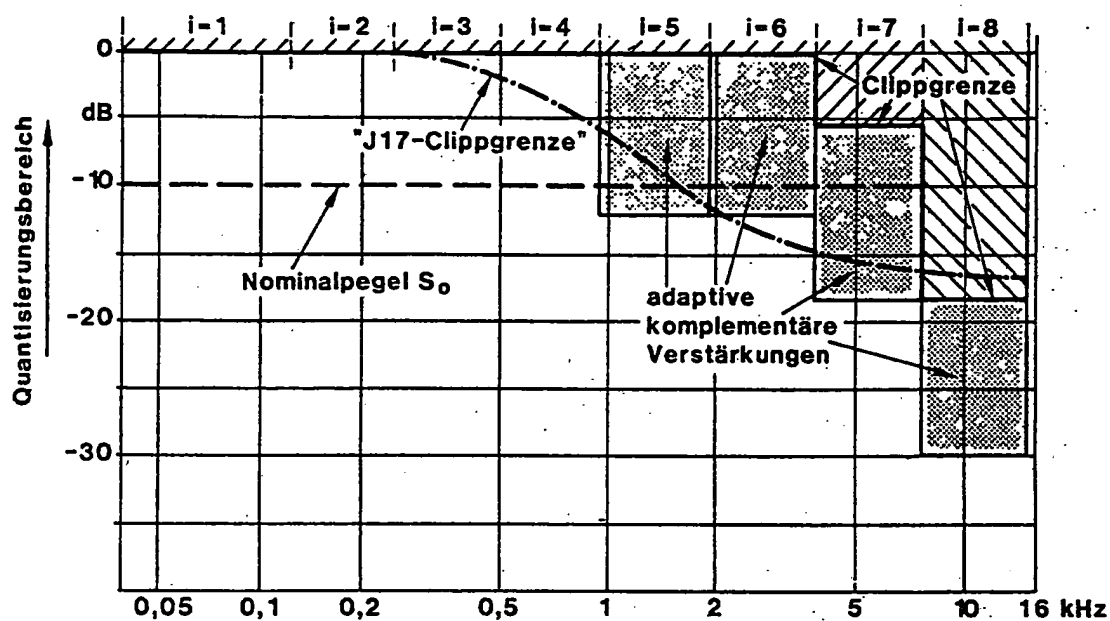
S/N-Diagramm zur Dimensionierung einer
 ohrangepaßten Bandpaß-PCM-Strecke

Fig. 8



Pre/Deemphasis einer Bandpaß-PCM-Strecke. Im Vergleich "J17-Clippgrenze" und Ruhehörschwelle

Fig. 9



Statische und adaptive komplementäre Verstärkungen in den Bandpässen einer Oktavband-PCM-Strecke. Im Vergleich "J-17 Clippgrenze"

Fig. 10